

Tilburg University

Versta ik het wel goed?

Vroomen, J.

Published in:
Psychologie

Publication date:
1996

[Link to publication in Tilburg University Research Portal](#)

Citation for published version (APA):
Vroomen, J. (1996). Versta ik het wel goed? *Psychologie*, (2), 24-25.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



versta ik het wel goed?

Wij denken vaak dat mensen die een ons onbekende taal spreken, tijdens de spraak geen adempauzes nemen en bovendien uitzonderlijk snel praten. Is dat ook zo?

JEAN VROOMEN

Het voeren van een gesprek is voor veel mensen de normaalste zaak van de wereld. Een gedachte in ons hoofd zetten we om in zinnen en woorden. Die woorden vertalen we op hun beurt in bewegingen die uiteindelijk de spraakklanken voortbrengen. Tijdens de overdracht van deze klanken naar de luisteraar, kunnen allerlei achtergrondgeluiden het spraaksignaal verstoren. Bovendien wordt de spraak vervormd door de akoestische eigenschappen van de ruimte waarin de spreker zich bevindt. Uiteindelijk zal het spraaksignaal het oor van de luisteraar bereiken die er, ondanks alle vervormingen, de woorden in zal herkennen die de spreker uitsprak. Op dat moment kan de luisteraar zich een begrip vormen van wat de spreker bedoelde en is de communicatie geslaagd. Dit is een zeer korte beschrijving van de manier waarop in een gesprek een gedachte van de ene persoon naar de ander wordt overgebracht. Een van de aspecten in dit proces dat reeds lang de aandacht van psychologen trekt, is hoe het mogelijk is dat woorden zo snel, zo moeiteloos en zo foutloos herkend kunnen worden, ondanks het feit dat er allerlei storingen (zoals achtergrondgeluiden) kunnen optreden. Wat in ieder geval duidelijk is, is dat spraakherkenning een automatisch proces is. Bij de mens is de herkenning van spraak zelfs zo automatisch, dat we nauwelijks idee hebben van de complexe processen die eraan ten grondslag liggen. Waarschijnlijk is dit maar goed ook, omdat we anders niet toe zouden komen aan het begrijpen van wat er gezegd wordt.

Spraakonderzoekers zijn zich bewust geworden van de enorme complexiteit van het spraaksignaal op het moment dat men computers wilde maken die spraak automatisch zouden kunnen herkennen. Voor een aantal toepassingen zou dit handig zijn. Je zou bijvoorbeeld het toetsenbord van een computer kunnen vervangen door een microfoon waardoor het type-diploma een relikwie zou worden. Ook kan men denken aan de vervanging van de druktoetsen van een telefoon door een microfoon. Men zou dan alleen tegen de telefoon hoeven te zeggen even naar een vriendin te bellen en aan de andere kant zou de telefoon reeds overgaan. Dergelijke toepassingen beginnen langzaam werkelijkheid te worden, maar dit betekent nog niet dat de fundamentele problemen in spraakherkenning zijn opgelost.

Een van de grootste problemen waar een automatische spraakherkenner mee te maken heeft, staat bekend als het *segmentatieprobleem*, het opdelen van klanken in woorden. De achtergrond van dit probleem is dat woorden in natuurlijke spraak bijna altijd ingebed zijn in een zin. In principe zou dit geen probleem hoeven te zijn, mits er tussen de afzonderlijke woorden merktekentjes aanwezig zouden zijn die de grenzen tussen woorden markeren. Men kan daarbij denken aan een geschreven zinnetje, waarin de spaties gebruikt worden om de grenzen tussen woorden aan te geven. In het spraaksignaal zijn echter geen hoorbare spaties aanwezig. Weliswaar kan in een aantal gevallen een pauze dienen als woordgrens, maar korte pauzes zijn vaak onbetrouwbaar. Een zin als 'de melk is zuur' bijvoorbeeld, wordt in lopende spraak uitgesproken als 'demél kisuur' (zie kader). Er is dus wel een pauze in die zin, maar die staat voor de /k/ binnen het woord 'melk', terwijl er juist geen pauze is tussen 'de melk' en 'is zuur'. Korte pauzes zijn om die reden nauwelijks betrouwbaar. Bij automatische spraakherkenning wordt het segmentatieprobleem vaak opgelost door de spreker te leren ieder woord afzonderlijk uit te spreken met grote pauzes tussen ieder woord. Dit is natuurlijk een onbevredigende oplossing omdat het ongemakkelijk is voor de spreker en omdat het segmentatieprobleem alleen wordt omzeild, maar niet werkelijk wordt opgelost.

WOORDENBRIJ Door het segmentatieprobleem weet een automatisch spraakherkenningssysteem vaak niet waar het ene woord begint en het andere eindigt. Intuïtief is dit voor te stellen door naar een vreemde taal als het Chinees of Japans te luisteren. Het begin of het einde van een hele zin kan nog wel met enige zekerheid worden vastgesteld, maar het begin of het einde van een woord is nauwelijks te ontdekken. Vaak lijken buitenlandse talen dan ook op een grote brij van klanken die bovendien ook nog veel te snel wordt uitgesproken.

Deze 'klankbrij-kwaliteit' is echter geen typisch Japans of Chinees fenomeen. Het blijkt dat veel talen, ook het Nederlands, een brij worden in de oren van een buitenlander. De zaak verandert echter op het moment dat een aantal woorden in de vreemde taal bekend is. Een typisch voorbeeld is hoe het Deens in de oren van een Nederlander klinkt. De meeste Deense woorden zijn ons onbekend, maar bij het luisteren naar het Deens, springt er soms als het ware een woord uit omdat dit wel erg veel lijkt op een Nederlands equivalent. Dit 'eruit springen' van woorden is nu precies waar het om draait. Op deze wijze worden woorden in de eigen moedertaal ook herkend – met dit verschil dat in de moedertaal *ieder* woord er uitspringt. De vraag is hoe we ons dit 'eruit springen' moeten voorstellen. Hoe kunnen we

een woord herkennen terwijl we niet eens weten waar het begint?

Bij de Nederlandse taal hebben we in onderzoek gekeken naar het onderscheid tussen sterke en zwakke lettergrepen. Een sterke lettergreep kan gedefinieerd worden als een lettergreep met een volle klinker, terwijl een zwakke een minder volle klinker heeft zoals de 'uh' die te horen is in de tweede klinker van het woord 'vader'. In het woord 'vader' is de eerste lettergreep sterk, de tweede dus zwak. Uit een analyse van de Nederlandse woordenschat blijkt dat woorden veel vaker beginnen met een sterke lettergreep dan met een zwakke. Het woord 'vader' is dus een typisch voorbeeld. (Een uitzondering op deze regelmatigheid zijn woorden die met 'be-', 'ge-' of 'ver-' beginnen, alhoewel dit vaak afgeleide woorden zijn zoals 'bedankt', 'geleefd' en 'verliefd'). Het idee van een zogenaamde *metrische segmentatie* is dat luisteraars sterke lettergrepen als 'merktekens' gebruiken om het begin van een nieuw woord te vermoeden. Dus, iedere keer als een sterke lettergreep wordt gehoord, wordt een nieuw woord opgezocht in het mentale woordenboek dat met deze lettergreep begint.

Deze metrische-segmentatieprocedure kan een aantal typische fouten in de spraakwaarneming verklaren. Om zulke fouten op een gecontroleerde manier uit te lokken, kunnen we spraak op een zeer laag geluidsniveau aanbieden zodat deze nog net verstaanbaar is (dit klinkt alsof de radio heel zacht staat). In ons onderzoek lieten we bijvoorbeeld korte onvolledige zinsneden horen die met veel moeite konden worden verstaan. Proefpersonen moesten vervolgens zeggen wat ze hadden gehoord. Een typisch voorbeeld van zo'n zinsnede is 'inTERN beSLUIT' welk een metrisch patroon heeft dat gelijk is aan 'Zwak-Sterk Zwak-Sterk'. Proefpersonen waren geneigd om dit voor het Nederlands a-typische patroon te veranderen in een regelmatig patroon waarin de woorden met een sterke klinker beginnen. De zin 'inTERN beSLUIT' werd bijvoorbeeld gehoord als 'de KERker SLUIT'. In dit voorbeeld is te zien dat er twee woordgrenzen zijn veranderd die beide in overeenstemming zijn met de principes van de metrische segmentatie. Een foutieve woordgrens werd ingelast vóór de sterke klinker 'TERN' ('inTERN' -> 'de KER') en er werd een woordgrens weggelaten vóór de zwakke klinker 'be' ('TERN be' -> 'KERker'). Beide typen van fouten kwamen vaker voor dan men op basis van kans alleen zou kunnen verwachten. Blijkbaar gebruiken wij bij het luisteren naar spraak deze statistische eigenschap van de taal. ●

